

# Large Deviations in Stochastic Heat-Conduction Processes Provide a Gradient-Flow Structure for Heat Conduction

Mark A. Peletier,<sup>1, a)</sup> Frank Redig,<sup>2, b)</sup> and Kiamars Vafayi<sup>3, c)</sup>

<sup>1)</sup>*Department of Mathematics and Computer Science and Institute for Complex Molecular Systems, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, The Netherlands*

<sup>2)</sup>*Delft Institute of Applied Mathematics, Technische Universiteit Delft, Mekelweg 4, 2628 CD Delft, The Netherlands*

<sup>3)</sup>*Department of Mathematics and Computer Science, Technische Universiteit Eindhoven, Postbus 513, 5600 MB Eindhoven, The Netherlands*

(Dated: 21 March 2014)

We consider three one-dimensional continuous-time Markov processes on a lattice, each of which models the conduction of heat: the family of Brownian Energy Processes with parameter  $m$ , a Generalized Brownian Energy Process, and the Kipnis-Marchioro-Presutti process. The hydrodynamic limit of each of these three processes is a parabolic equation, the linear heat equation in the case of the BEP( $m$ ) and the KMP, and a nonlinear heat equation for the GBEP( $a$ ). We prove the hydrodynamic limit rigorously for the BEP( $m$ ), and give a formal derivation for the GBEP( $a$ ).

We then formally derive the pathwise large-deviation rate functional for the empirical measure of the three processes. These rate functionals imply gradient-flow structures for the limiting linear and nonlinear heat equations. We contrast these gradient-flow structures with those for processes describing the diffusion of mass, most importantly the class of Wasserstein gradient-flow systems. The linear and nonlinear heat-equation gradient-flow structures are each driven by entropy terms of the form  $-\log \rho$ ; they involve dissipation or mobility terms of order  $\rho^2$  for the linear heat equation, and a nonlinear function of  $\rho$  for the nonlinear heat equation.

---

<sup>a)</sup>Electronic mail: [m.a.peletier@tue.nl](mailto:m.a.peletier@tue.nl)

<sup>b)</sup>Electronic mail: [f.h.j.redig@tudelft.nl](mailto:f.h.j.redig@tudelft.nl)

<sup>c)</sup>Electronic mail: [k.vafayi@tue.nl](mailto:k.vafayi@tue.nl)

## I. THE HEAT EQUATION

The equation

$$\partial_t \rho = \Delta \rho \tag{1}$$

is known both as the ‘diffusion equation’ and as the ‘heat equation’. The two names reflect the fact that one arrives at the same equation via two completely different modeling routes—one describing the motion of particles, the other the conduction of heat.

Recently, one of these modeling routes has received special attention. Inspired by the observation that equation (1) is a Wasserstein gradient flow (we describe this below), and that the Wasserstein gradient flows have many interesting and useful properties, we asked the question ‘how is the Wasserstein gradient-flow structure of (1) related to the modeling of (1)?’ We answered this question<sup>1–3</sup> by connecting the Wasserstein gradient-flow structure to the large-deviation behavior of stochastic particle systems.

Before turning to the content of the present paper, we briefly describe these previous results. A *gradient flow* on a linear space  $X$  is formally defined by a functional  $E : X \rightarrow \mathbb{R}$  and a symmetric operator  $K : X' \rightarrow X$  (the so-called *Onsager operator*) where  $X'$  denotes the dual space of  $X$ . The equation (1) has many such gradient-flow structures, each corresponding to a different triple  $(X, E, K)$ . One example is  $X = X' = L^2(\mathbb{R}^n)$ ,  $E(\rho) = \frac{1}{2} \int_{\mathbb{R}^n} |\nabla \rho|^2$ , and  $K\xi = \xi$ ; then by definition

$$\partial_t \rho = -K \left( \frac{\delta E}{\delta \rho}(\rho) \right) = -K(-\Delta \rho) = \Delta \rho.$$

Note that we specify  $K$  by describing its effect on the variational derivative  $\delta E / \delta \rho$  of  $E$ . Other examples are

$$X = H^{-1}(\mathbb{R}^n), \quad E(\rho) = \frac{1}{2} \int_{\mathbb{R}^n} \rho^2, \quad K\xi = -\Delta \xi,$$

$$X = L^1_{\geq 0}(\mathbb{R}^n), \quad E(\rho) = \int_{\mathbb{R}^n} \rho \log \rho, \quad K_\rho \xi = -\operatorname{div} \rho \nabla \xi, \tag{2}$$

$$X = L^1(\mathbb{R}^n), \quad E(\rho) = \int_{\mathbb{R}^n} f(\rho), \quad K_\rho \xi = -\operatorname{div} \left( \frac{1}{f'(\rho)} \nabla \xi \right). \tag{3}$$

Each of these gradient-flow structures leads to the same equation (1). The last two examples show that  $K$  may depend on  $\rho \in X$ ; in geometric terms,  $K$  maps the cotangent bundle to the tangent bundle, and therefore may depend on position.

Example (2) is the Wasserstein gradient-flow structure that was mentioned above. We showed in 1 and 2 how the Wasserstein structures of this equation and of more general

convection-diffusion equations arise from the properties of an underlying stochastic process: systems of particles whose positions are described by a stochastic differential equation. The large deviations of empirical measures of these particles directly identify a gradient-flow structure, which is of the form (2). This identification consists of two parts:

1. The equilibrium large-deviation rate functional of the *invariant measure* is the driving functional  $E$ ;
2. The non-equilibrium large-deviation rate functional of the *law of the time-courses of the empirical measure* identifies the operator  $K$ .

Through these identifications, the properties of this gradient-flow structure can be traced back to the properties of the underlying stochastic process. For instance, the presence or absence of interaction only appears in the form of the driving functional  $E$ , and the form of  $K$  is completely determined by the Brownian noise in the SDE. As a result, the Wasserstein gradient-flow structure of (1) can now be completely explained in terms of properties of such underlying models of diffusion. The same arguments apply to various related equations<sup>4–7</sup>, and we have recently showed how it is shared among all processes with detailed balance<sup>8</sup>. The insight that this provides is of significant use in the modeling of other real-world systems.

However, this explanation only applies to *diffusion*—to the derivation of (1) in which  $\rho$  is a *density of particles* that evolves by thermal agitation. For the other interpretation, where  $\rho$  is a *temperature* or a *heat content*, and where heat disperses through heat conduction, very little is known in this direction. Is there a corresponding gradient-flow structure of (1), that derives from large deviations of some stochastic process? If so, is it the same as for diffusion, or a different one?

In this paper we investigate this question. We study stochastic processes that mimic the *conduction of heat*: in these processes a local ‘internal’ energy is exchanged through random nearest-neighbor interactions. The hydrodynamic or continuum limit of these systems is again a parabolic partial differential equation—and in some of the cases it is exactly the linear heat equation (1). As in the case of diffusion, the large deviations away from this continuum limit identify a natural gradient-flow structure, that we discuss in detail below.

In fact, we will consider three stochastic processes:

1. The *Brownian Energy Process* with parameter  $m$  (BEP( $m$ ), where  $m$  is a positive real number);

2. A *Generalized* Brownian Energy Process with parameter  $a$  (GBEP( $a$ ), where  $a$  is a positive function);
3. The *Kipnis-Marchioro-Presutti* process (KMP).

Each of these describes heat conduction in a slightly different way, and we will use the similarities and differences to investigate how different aspects of these stochastic processes lead to different continuum limits, and to different corresponding gradient-flow structures. This will allow us to link the properties of the gradient-flow ingredients to specific aspects of the underlying stochastic processes.

The outline of this paper is as follows. In Section II we describe the three processes mentioned above. Most of the further development is done for the BEP( $m$ ): in Sections III and IV we study the equilibrium properties of the BEP, and in Sections V, VI, and VIII the hydrodynamic limit and the corresponding large deviation principle. In Section IX we discuss how the corresponding properties of the two other processes can be derived. Finally, in Section X we discuss the gradient-flow structures that these large deviations suggest, and in Section XI the implications of these structures.

## II. THE THREE STOCHASTIC PROCESSES: BEP( $m$ ), GBEP( $a$ ), AND KMP

Heat conduction is the transfer of internal energy between neighboring parts of a material. In the most straightforward microscopic models of heat conduction, matter is represented as a collection of oscillators on lattice points; each oscillator is a little isolated Hamiltonian system, such as a little mass-spring system. If there is no interaction, then the oscillator has a finite energy—the Hamiltonian—that is conserved over time.

Heat conduction arises when one introduces interaction between the oscillators, either deterministically or stochastically, allowing energy exchange between neighboring sites. Deterministic interaction between neighboring oscillators is natural, since both classical and quantum descriptions of reality are Hamiltonian; however, the difficulties surrounding the analysis of deterministic heat-conducting systems are formidable (see e.g. 9 and 10). For this reason the interaction is often chosen to be stochastic, and many varieties exist in addition to the ones we study here; see e.g. 11–18.

The models that we study in this paper are all relatives of a single, prototypical model, that we now describe. Afterwards, we introduce the three models which are the subject of this paper and explain how they are related.

## A. The Brownian Momentum Process

In the *Brownian Momentum Process*, physical space is represented by a one-dimensional lattice of  $N$  points  $i = 1, \dots, N$ ; we silently assume periodicity modulo  $N$ , i.e. the point  $i = N + 1$  is the same as the point  $i = 1$ . The *state* of the BMP process is a vector of *momenta*  $p_i$ ,  $i = 1, \dots, N$ . The BMP process is characterized by its generator  $L^{\text{BMP}}$  that acts on smooth functions  $f : \mathbb{R}^N \rightarrow \mathbb{R}$ , given by

$$L^{\text{BMP}} = \sum_{i=1}^N L_{i,i+1}^{\text{BMP}}, \quad L_{i,i+1}^{\text{BMP}} f = \frac{1}{2} \left( p_i \frac{\partial}{\partial p_{i+1}} - p_{i+1} \frac{\partial}{\partial p_i} \right)^2 f. \quad (4)$$

Equivalently, the  $p_i$  solve the Itô stochastic differential equation

$$dp_i = p_{i+1} dB_{i,i+1} - p_{i-1} dB_{i-1,i} - p_i dt, \quad (5)$$

where the  $B_{i,i+1}$  are independent Brownian motions.

The pairwise generator  $L_{i,i+1}^{\text{BMP}}$  can be interpreted as a Brownian motion on the circle given by the condition  $p_i^2 + p_{i+1}^2 = \text{constant}$ . This follows from noting that  $L_{i,i+1}^{\text{BMP}}$  preserves  $p_i^2 + p_{i+1}^2$ , and in terms of the angle coordinate  $\varphi = \arctan(p_{i+1}/p_i)$  we can write  $L_{i,i+1}^{\text{BMP}} = \frac{1}{2} \partial^2 / \partial \varphi^2$ , representing a simple Brownian motion in  $\varphi$ . In terms of the SDE (5), the conservation of  $p_i^2 + p_{i+1}^2$  can be observed in the double occurrence of the noise  $dB_{i,i+1}$ , once in the equation for  $p_i$  (with prefactor  $p_{i+1}$ ) and once in the equation for  $p_{i+1}$ , with the opposite sign, and with prefactor  $p_i$ . The term  $-p_i dt$  is an Itô correction that compensates for the asymmetry in the definition of an Itô SDE.

This process is a model of heat conduction in the following sense. Up to a constant, the sum  $p_i^2 + p_{i+1}^2$  should be interpreted as the total kinetic energy of the two sites  $i$  and  $i + 1$ . Therefore each neighbour interaction, characterized by  $L_{i,i+1}^{\text{BMP}}$ , is a Brownian motion over all possible distributions of the preserved kinetic energy over the two momenta  $p_i$  and  $p_{i+1}$ , and therefore implements a stochastic exchange of energy between neighbouring sites.

## B. The Brownian Energy Process with parameter $m$

Because of our focus on heat conduction, we shall be interested in tracking the squares  $z_i := p_i^2$ , i.e. the site energies, rather than in the momenta  $p_i$  themselves. The *Brownian Energy Process* (BEP), with parameter *one*, is the process that governs the evolution of the vector of energies  $(z_1, \dots, z_N)$ . It happens to be a Markov process itself: the evolution of the  $z_i$  can be characterized in terms of the  $z_i$  themselves, with generator

$$L^{\text{BEP}} = \sum_{i=1}^N L_{i,i+1}^{\text{BEP}},$$

where

$$L_{i,i+1}^{\text{BEP}} = 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 - (z_i - z_{i+1}) (\partial_i - \partial_{i+1}). \quad (6)$$

Here we write  $\partial_i := \partial/\partial z_i$  for the derivative with respect to the variable  $z_i$ . The corresponding SDE is

$$dz_i = 2\sqrt{z_i z_{i+1}} dB_{i,i+1} - 2\sqrt{z_{i-1} z_i} dB_{i-1,i} + (z_{i+1} - 2z_i + z_{i-1}) dt. \quad (7)$$

This process is the Brownian Energy Process—with parameter one.

The *Brownian Energy Process with parameter  $m$* ,  $\text{BEP}(m)$ , is a generalization of this process, in which not a single scalar quantity  $p_i$ , but  $m$  scalar quantities  $p_i^j$ ,  $j = 1, \dots, m$  are defined at each lattice point  $i$ . We define the generator for this  $m$ -fold BMP by imposing the same interaction between all  $m \times m$  neighboring pairs, i.e.

$$L^{\text{BMP}(m)} = \sum_{i=1}^N L_{i,i+1}^{\text{BMP}(m)}, \quad L_{i,i+1}^{\text{BMP}(m)} = \sum_{j,j'=1}^m \frac{1}{2} \left( p_i^j \frac{\partial}{\partial p_{i+1}^{j'}} - p_{i+1}^{j'} \frac{\partial}{\partial p_i^j} \right)^2.$$

The total energy for each site  $i$  is now defined as  $z_i := \sum_{j=1}^m (p_i^j)^2$ , and this is again a Markov process with generator

$$L^{\text{BEP}(m)} = \sum_{i=1}^N L_{i,i+1}^{\text{BEP}(m)}, \quad L_{i,i+1}^{\text{BEP}(m)} = 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 - m(z_i - z_{i+1}) (\partial_i - \partial_{i+1}). \quad (8)$$

Note the small difference with the generator (6) of the  $\text{BEP}(1)$ : only the coefficient of the drift term is  $m$ -dependent. Also note that for the definition of the process  $z_i$ ,  $m$  need not be integer—any  $m > 0$  is admissible, although the connection to an underlying Momentum Process of course only exists for integer  $m$ .

The Brownian Momentum Process was introduced in 19 with imposed-temperature boundary conditions, and further studied in 20–22. In 22 we studied non-equilibrium stationary states of this system, generated by imposing different temperatures at the two ends of a one-dimensional system. We showed, among other things, that the expectation of the temperature along these non-equilibrium steady states is linear. This is of course consistent with non-equilibrium steady states of the macroscopic equation (1). The Brownian Energy Process was introduced in 20 and further studied in 21 and 23.

### C. A Generalized Brownian Energy Process with parameter $a$

For reasons that will become clear below, it is interesting to consider a different generalization of the BEP, in which the noise terms in (5) and (7) are modified in an energy-conserving manner, as follows. Choose a continuous function  $a : [0, \infty) \rightarrow (0, \infty)$ , and define

$$a_{i,i+1} := a\left(\frac{1}{2}(z_i + z_{i+1})\right).$$

The *Generalized Brownian Momentum Process* is the SDE

$$dp_i = p_{i+1}a_{i,i+1} dB_{i,i+1} - p_{i-1}a_{i-1,i} dB_{i-1,i} - \frac{1}{2}p_i(a_{i,i+1}^2 + a_{i-1,i}^2) dt, \quad (9)$$

and the corresponding *Generalized Brownian Energy Process* with parameter  $a$  is the SDE

$$dz_i = 2a_{i,i+1}\sqrt{z_i z_{i+1}} dB_{i,i+1} - 2a_{i-1,i}\sqrt{z_{i-1} z_i} dB_{i-1,i} + ((z_{i+1} - z_i)a_{i,i+1}^2 + (z_i - z_{i-1})a_{i-1,i}^2) dt, \quad (10)$$

with generator

$$L^{\text{GBEP}(a)} := \sum_{i=1}^N a_{i,i+1}^2 \left[ 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 + (z_{i+1} - z_i)(\partial_i - \partial_{i+1}) \right] = \sum_{i=1}^N a_{i,i+1}^2 L_{i,i+1}^{\text{BEP}}.$$

As far as we are aware, this process has not been discussed in the literature.

### D. The Kipnis-Marchioro-Presutti process

In 11 Kipnis, Marchioro, and Presutti introduced a model for heat conduction. In the KMP process energy is not exchanged continuously, as in the processes described above, but discretely: at exponentially distributed times we choose a random pair of neighbors,

say with energies  $z_i$  and  $z_{i+1}$ ; we then equidistribute the joint energy  $z_i + z_{i+1}$  over the two sites, with a fraction  $s$  on one site and a fraction  $1 - s$  on the site. Here  $s$  is drawn from the uniform distribution on  $[0, 1]$ . The KMP process has the generator<sup>24</sup>

$$(L^{\text{KMP}} f)(z) := 2 \sum_{i=1}^N \int_0^1 [f(z^{i,i+1,s}) - f(z)] ds,$$

where the modified  $z$  is defined as

$$z_j^{i,i+1,s} := \begin{cases} s(z_i + z_{i+1}) & \text{if } j = i, \\ (1-s)(z_i + z_{i+1}) & \text{if } j = i+1, \\ z_j & \text{otherwise.} \end{cases}$$

As pointed out in 20, the KMP process is related to the process BEP(2), in the following way. Choose any  $i$  and consider the single-pair generator  $L_{i,i+1}^{\text{BEP}(m)}$ . This single-pair generator has a unique invariant measure for every  $m$ , which describes the equilibrium distribution of energy over the two sites  $i$  and  $i+1$ , assuming all other  $z_j$  are fixed. For  $m = 2$ , this measure is exactly the uniform distribution of energy between the two sites  $i$  and  $i+1$ . (The general case is that if  $(z_i, z_{i+1}) = (E - e, E + e)$ , then  $e \in [0, E]$  has probability density  $\sim (E^2 - e^2)^{m/2-1}$ .)

Therefore  $L^{\text{KMP}}$  can be considered an ‘instantaneously equilibrated’ or ‘instantaneously thermalised’<sup>20</sup> version of  $L^{\text{BEP}(2)}$ —as if the exchange between  $i$  and  $i+1$  is not smeared out over time but concentrated into very short bursts, during which, complete equilibration within that specific bond takes place.

## E. Macroscopic quantities

In each of the three processes BEP( $m$ ), GBEP( $a$ ), and KMP, the variable  $z_i$  at site  $i$  represents the local or microscopic energy at that site. The natural way to connect this local energy with macroscopic quantities such as temperature and macroscopic energy is as follows. Given an energy state  $Z = \{z_1, z_2, \dots, z_N\}$ , we define the empirical energy measure  $\pi_N(Z)$  to be the probability measure

$$[\pi_N(Z)](dx) := \frac{1}{N} \sum_{i=1}^N z_i \delta_{i/N}(dx), \quad (11)$$



where  $\delta_{i/N}(x)$  is the Delta measure at  $x = i/N$ . This makes  $\pi_N$  a probability measure on the flat one-dimensional torus  $\mathbb{T} = \mathbb{R}/\mathbb{Z}$ , i.e. the interval  $[0, 1]$  with periodic boundary conditions.

We show below how, in the limit  $N \rightarrow \infty$ ,  $\pi_N$  converges weakly to a deterministic limit profile  $\rho$ , both in equilibrium and in non-equilibrium situations. The limit  $\rho$  will be interpreted as macroscopic energy. The large deviations away from the deterministic limit will provide us with the structure of the corresponding gradient flow.

### III. EQUILIBRIUM PROPERTIES OF THE BEP( $m$ )

In the next few sections we focus on the BEP( $m$ ); we study the hydrodynamic limit and the large deviations of the model, and we derive from the large deviations the corresponding gradient-flow structure. In Section IX we describe the differences for the two other processes.

We first consider equilibrium properties. Since  $L^{\text{BEP}(m)}[\sum_i z_i] = 0$ , the BEP( $m$ ) deterministically conserves the total energy  $\sum_i z_i$ —this property is also natural from the construction of the process. As a consequence the process does not have a single invariant measure, but a large collection of invariant measures, generated by a particular one-parameter family of invariant measures, indexed by a parameter  $\theta > 0$ . For each  $\theta$ , with the interpretation of temperature, this invariant measure is a product measure over the sites  $i$ :

$$\nu_{N,\theta,m}(dZ) = \prod_{i=1}^N \nu_{\theta,m}(dz_i), \quad \text{where} \quad \nu_{\theta,m}(dz) = \frac{1}{\theta^{m/2} \Gamma\left(\frac{m}{2}\right)} z^{-1+m/2} e^{-z/\theta} dz, \quad (12)$$

and  $\Gamma$  is the Gamma-function

$$\Gamma(k) := \int_0^\infty e^{-t} t^{k-1} dt.$$

The process is reversible with respect to each of these invariant measures.

In equilibrium, since this invariant measure is a product measure, to every parameter  $\theta$  corresponds a uniform (i.e.,  $i$ -independent) average energy density  $\rho_0$  given by

$$\rho_0(\theta) = \mathbb{E}_{\nu_{\theta,m}}[z_i] = \frac{1}{\theta^{m/2} \Gamma\left(\frac{m}{2}\right)} \int_0^\infty z^{m/2} e^{-z/\theta} dz = \frac{m\theta}{2}. \quad (13)$$

Since there is a one-to-one correspondence between  $\theta$  and  $\rho_0$ , we can equivalently index the empirical measure by  $\rho_0$  as  $\nu_{N,\rho_0,m} = \nu_{N,\theta(\rho_0),m}$  instead.

Moreover, if  $Z$  is distributed according to  $\nu_{N,\rho_0,m}$ , the empirical energy  $\pi_N(Z)$  converges in probability to the constant measure  $\rho_0$ , by the law of large numbers: for each  $\delta > 0$  and each test function  $\phi \in C_b(\mathbb{T})$  we have

$$\lim_{N \rightarrow \infty} \nu_{N,\rho_0,m} \left( |\langle \pi_N(Z), \phi \rangle - \langle \rho_0, \phi \rangle| > \delta \right) = 0 \quad (14)$$

where

$$\langle \rho_0, \phi \rangle := \int_{\mathbb{T}} \rho_0(x) \phi(x) \, dx.$$

#### IV. EQUILIBRIUM LARGE DEVIATIONS OF THE BEP

We already mentioned in the introduction that for reversible processes, such as the three processes of this paper, the large-deviations rate functional of the invariant measure is the functional that drives the gradient-flow evolution<sup>8</sup>. We now calculate this rate functional.

A large-deviation principle states that the probability that the empirical energy  $\pi_N(Z)$  deviates considerably from its most likely value  $\rho_0$  and thus becomes close to some other profile  $\rho$  is exponentially small in  $N$  for large  $N$  and is determined by a rate functional  $S_{\rho_0,m}$  according to

$$\nu_{N,\rho_0,m}(\pi_N(Z) \approx \rho) \sim \exp\{-N S_{\rho_0,m}(\rho)\}.$$

More precisely,

**Theorem 1.** *For all Borel subsets  $A$  of the set of non-negative measures  $\mathcal{M}_+(\mathbb{T})$ ,*

$$-\inf_{\text{int } A} S_{\rho_0,m} \leq \liminf_{N \rightarrow \infty} \frac{1}{N} \log \nu_{N,\rho_0,m}(A) \leq \limsup_{N \rightarrow \infty} \frac{1}{N} \log \nu_{N,\rho_0,m}(A) \leq -\inf_{\text{cl } A} S_{\rho_0,m}, \quad (15)$$

where

$$S_{\rho_0,m}(\rho) = \frac{m}{2} \int_0^1 \left( \frac{\rho(x)}{\rho_0} - 1 - \log \frac{\rho(x)}{\rho_0} \right) dx. \quad (16)$$

The existence of a functional  $S_{\rho_0,m}$  satisfying (15) and its explicit form both follow from the Gärtner-Ellis theorem 25, Th. 4.5.20; the proof is standard, and we sketch only those details that we need later on.

*Proof.* We first calculate  $S_{\rho_0,m}$ . Set  $\theta = 2\rho_0/m$  (following (13)) and define the logarithmic cumulant-generating function as

$$G : C_b(\mathbb{T}) \rightarrow \mathbb{R}, \quad G(\phi) := \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}_{\nu_{N,\theta,m}} \left( e^{N \langle \phi, \pi_N(Z) \rangle} \right). \quad (17)$$

From the definition of the energy profile we have, for continuous  $\phi \in C_b(\mathbb{T})$ ,

$$N\langle\phi, \pi_N(Z)\rangle = \sum_{i=1}^N \phi_i z_i, \quad \text{where we write } \phi_i := \phi(i/N).$$

Due to the independence of different sites in the product measure we can use the single-site generating function

$$\mathbb{E}_{\nu_{\theta,m}}[e^{\phi_i z_i}] = \frac{1}{\theta^{m/2} \Gamma(\frac{m}{2})} \int_0^\infty z^{-1+m/2} e^{-z(-\phi_i+1/\theta)} dz = \left( \frac{1}{1-\theta\phi_i} \right)^{m/2} \quad (18)$$

to calculate

$$\begin{aligned} G(\phi) &= \lim_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{E}_{\nu_{N,\theta,m}} \left( \prod_i e^{\phi_i z_i} \right) \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log \left( \prod_i \mathbb{E}_{\nu_{\theta,m}} e^{\phi_i z_i} \right) \quad \text{by independence} \\ &= \lim_{N \rightarrow \infty} \frac{1}{N} \log \prod_i \left( \frac{1}{1-\theta\phi_i} \right)^{m/2} \quad \text{by (18)} \\ &= -\frac{m}{2} \lim_{N \rightarrow \infty} \frac{1}{N} \sum_i \log(1-\theta\phi_i) \\ &= -\frac{m}{2} \int_{\mathbb{T}} \log(1-\theta\phi(x)) dx. \end{aligned}$$

(Note how the local nature of the functional  $G$  (the fact that  $G$  is a simple integral over  $\mathbb{T}$ ) arises from the product structure of the measure  $\nu_{N,\theta,m}$ .) We then obtain the formula (16) for  $S_{\rho_0,m}$  as the Legendre transform of  $G$  by a direct calculation.

After this calculation the proof of the property (15) follows from 25, Th. 4.5.20(c). There is one non-trivial condition to be verified, which is the density of an appropriate class of exposed points of  $\mathcal{M}_+(\mathbb{T})$ ; this can be ascertained by noting that the subset of probability measures with Lebesgue density that is continuous and bounded away from zero is dense in  $\mathcal{M}_+(\mathbb{T})$  with the narrow topology and satisfies the remaining requirements.  $\square$

## V. HYDRODYNAMIC LIMIT OF THE TIME EVOLUTION

We now turn to the time evolution of the BEP( $m$ ). First we note that the natural time scale for the empirical measure  $\pi_N(Z)$  is  $N^2$ , which leads us to define the rescaled (accelerated) generator

$$L_N^{\text{BEP}(m)} := N^2 L^{\text{BEP}(m)}.$$

Next we describe the initial state. Choose a fixed measure  $\rho^0 \in \mathcal{M}_+(\mathbb{T})$ , and choose a sequence of vectors  $Z_N^0 = (z_{N,1}^0, \dots, z_{N,N}^0) \in \mathbb{R}^N$  such that  $\pi_N(Z_N^0) \rightharpoonup \rho^0$  as  $N \rightarrow \infty$ . Moreover we assume that  $z_{N,i}^0$  are bounded uniformly in  $N$ . By  $\mathbb{P}_N^{\text{BEP}(m)}$  we denote the distribution of the BEP( $m$ ) process, accelerated by a factor  $N^2$ —i.e. with generator  $L_N^{\text{BEP}(m)}$ —where the initial condition is equal to  $Z_N^0$ .  $\mathbb{E}_N^{\text{BEP}(m)}$  is the expectation with respect to  $\mathbb{P}_N^{\text{BEP}(m)}$ . From now on we also consider the process  $Z$  itself to be accelerated by a factor  $N^2$ , i.e.  $Z$  has law  $\mathbb{P}_N^{\text{BEP}(m)}$ .

The limit of the process will be the unique solution  $\rho$  of the linear heat equation,

$$\partial_t \rho(t, x) = m \partial_{xx} \rho(t, x), \quad \text{for } x \in \mathbb{T} \text{ and } t > 0. \quad (19)$$

with initial datum

$$\rho(0, \cdot) = \rho^0.$$

The following theorem specifies the exact sense in which  $\pi_N(Z)$  converges to  $\rho$ .

**Theorem 2.** *For all  $\phi \in C_b(\mathbb{T})$  and  $t > 0$ ,*

$$\lim_{N \rightarrow \infty} \mathbb{P}_N^{\text{BEP}(m)} \left( \left| \langle \pi_N(Z(t)), \phi \rangle - \langle \rho(t), \phi \rangle \right| > \delta \right) = 0. \quad (20)$$

*Also a stronger property holds: the whole random trajectory  $\{\pi_N(Z(t)) : 0 \leq t \leq T\}$  converges weakly to the Dirac measure concentrated on the trajectory  $\{\rho(x, t) dx : 0 \leq t \leq T\}$ , in the Skorokhod topology of  $D([0, T]; \mathcal{M}_+(\mathbb{T}))$ . Here  $\mathcal{M}_+(\mathbb{T})$  is the space of non-negative measures on  $\mathbb{T}$  with the topology of weak convergence.*

A sketch of proof is as follows. If  $\phi$  is a smooth function on  $\mathbb{T}$ , we have from the definition of the generator of Markov processes

$$\frac{d}{dt} \mathbb{E}_N^{\text{BEP}(m)} [\langle \pi_N(Z(t)), \phi \rangle] = \mathbb{E}_N^{\text{BEP}(m)} (L_N^{\text{BEP}(m)} \langle \pi_N(Z(t)), \phi \rangle).$$

From the definition of empirical energy we have

$$\langle \pi_N(Z(t)), \phi \rangle = \frac{1}{N} \sum_{i=1}^N z_i(t) \phi_i,$$

where again we write  $\phi_i := \phi(i/N)$ . For fixed  $i$ , let  $f_i$  be the coordinate function  $f_i(Z) := z_i$ . Then

$$\begin{aligned} L^{\text{BEP}(m)} f_i &= (L_{i-1,i}^{\text{BEP}(m)} + L_{i,i+1}^{\text{BEP}(m)}) z_i = \left[ 2z_{i-1} z_i (\partial_{i-1} - \partial_i)^2 - m(z_{i-1} - z_i) (\partial_{i-1} - \partial_i) \right. \\ &\quad \left. + 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 - m(z_i - z_{i+1}) (\partial_i - \partial_{i+1}) \right] z_i \\ &= m[z_{i-1} - 2z_i + z_{i+1}], \end{aligned}$$

so that

$$L_N^{\text{BEP}(m)} \langle \pi_N(Z(t)), \phi \rangle = N \sum_{i=1}^N (L^{\text{BEP}(m)} f_i)(Z(t)) \phi_i = mN \sum_{i=1}^N \phi_i [z_{i-1} - 2z_i + z_{i+1}].$$

We now apply summation by parts, using the periodicity of the lattice, to obtain

$$\begin{aligned} L_N^{\text{BEP}(m)} \langle \pi_N(Z(t)), \phi \rangle &= mN \sum_{i=1}^N z_i [\phi_{i-1} - 2\phi_i + \phi_{i+1}] \\ &= \frac{m}{N} \sum_{i=1}^N z_i \Delta_N \phi_i \approx m \langle \pi_N(Z(t)), \partial_{xx} \phi \rangle, \end{aligned}$$

where  $(\Delta_N \phi)_i := N^2 [\phi_{i-1} - 2\phi_i + \phi_{i+1}]$  is a discrete approximation of the Laplacian and in the last step we used the smoothness of  $\phi$ . Then

$$\frac{d}{dt} \langle \pi_N(Z(t)), \phi \rangle \approx m \langle \pi_N(Z(t)), \partial_{xx} \phi \rangle,$$

which is a weak formulation of (19).

The rigorous proof of Theorem 2 follows from standard arguments for gradient processes that can be found in e.g. 26, Ch. 4. These revolve around the martingale

$$\begin{aligned} M_t^N &:= \langle \pi_N(Z(t)), \phi \rangle - \langle \pi_N(Z_N^0), \phi \rangle - \int_0^t L_N^{\text{BEP}(m)} \langle \pi_N(\cdot), \phi \rangle (Z(s)) \, ds \\ &= \langle \pi_N(Z(t)), \phi \rangle - \langle \pi_N(Z_N^0), \phi \rangle - \int_0^t \langle \pi_N(Z(s)), m \partial_{xx} \phi \rangle \, ds + o(1) \end{aligned} \quad (21)$$

where  $o(1)$  converges to zero uniformly as  $N \rightarrow \infty$ .

The essential property of  $M^N$  that provides both compactness and convergence is the vanishing of the quadratic variation, the content of the following lemma.

**Lemma 3.** *The quadratic variation of  $M^N$  satisfies*

$$\forall t > 0, \quad \lim_{N \rightarrow \infty} \mathbb{E}_N^{\text{BEP}(m)} [M^N]_t = 0. \quad (22)$$

*Proof.* The quadratic variation of  $M_t^N$  is given by

$$[M^N]_t = \int_0^t \left( L_N^{\text{BEP}(m)}(g_N^2)(Z(s)) - 2g_N(Z(s)) L_N^{\text{BEP}(m)}(g_N)(Z(s)) \right) \, ds$$

where

$$g_N(Z(t)) = \frac{1}{N} \sum_{i=1}^N z_i(t) \phi_i.$$

Explicit computation then gives

$$\langle M_t^N \rangle = \frac{4}{N^2} \sum_{i=1}^N \phi_i^2 \int_0^t u_i(Z(s)) ds$$

with

$$u_i(z) = z_i z_{i+1} + z_{i-1} z_i.$$

We use the assumption that the initial condition is bounded, i.e., that  $\sup_i z_{N,i}^0 < C$ . Then we have

$$\mathbb{E}_N^{\text{BEP}(m)} \left( \sum_{i=1}^N \int_0^t \phi_i^2 u_i(Z(s)) ds \right) = \sum_{i=1}^N \phi_i^2 \int_0^t \mathbb{E}_N^{\text{BEP}(m)}(u_i(Z(s))) ds.$$

Therefore, in order to obtain (22), it suffices to see that

$$\mathbb{E}_N^{\text{BEP}(m)}(z_i(s) z_{i+1}(s))$$

is bounded uniformly in  $N$  and in  $0 \leq s \leq t$ . The expression  $z_i z_{i+1}$  is equal to  $K_m D(\delta_i + \delta_{i+1}, z)$ , where  $D(\xi, z)$  are the duality polynomials introduced in 19 and 21, and where  $K_m$  is a constant only depending on  $m$ . Therefore, by duality of the  $\text{BEP}(m)$  and the Symmetric Inclusion Process (SIP, see 19 and 21 for more details on this process and the duality), we have

$$\mathbb{E}_N^{\text{BEP}(m)}(z_i(s) z_{i+1}(s)) = K_m \mathbb{E}_{i,i+1}^{\text{SIP}} D(\delta_{X_t} + \delta_{Y_t}, z) \leq B_m(C + C^2),$$

where  $\mathbb{E}_{i,i+1}^{\text{SIP}}$  denotes expectation in the SIP starting with a single particle at both the sites  $i$  and  $i+1$ , and where in the last step we used that  $D(\delta_{X_t} + \delta_{Y_t}, z)$  is a second order polynomial in the variables  $z_i$ , with coefficients only depending on  $m$ .  $\square$

From (22), using Doob's inequality, we obtain that for all test functions  $\phi$

$$\sup_{0 \leq t \leq T} \left| \langle \pi_N(Z(t)), \phi \rangle - \langle \pi_N(Z_N^0), \phi \rangle - \int_0^t \langle \pi_N(Z(s)), m \partial_{xx} \phi \rangle ds \right|$$

converges to zero in probability, which is exactly the statement of the hydrodynamic limit in the path space version.

## VI. THE WEAKLY ASYMMETRIC $\text{BEP}(m)$

For the calculation of the nonequilibrium large deviations of the  $\text{BEP}(m)$  we need to calculate the hydrodynamic limit of a weakly-perturbed version of the  $\text{BEP}(m)$ , the Weakly

Asymmetric BEP( $m$ ) or WABEP( $m$ ):

$$L_N^{\text{WABEP}(m)} := L_N^{\text{BEP}(m)} + \Gamma.$$

There are various ways to introduce this asymmetry. From group-theoretic considerations in 19 and 21 we choose  $\Gamma$  to be

$$\Gamma = -N^2 \sum_{i=1}^{N-1} E_i z_i z_{i+1} (\partial_i - \partial_{i+1}),$$

where the  $E_i$  are to scale as  $1/N$  as  $N \rightarrow \infty$ . One way to achieve this, which will also simplify formulas later, is to choose

$$E_i := H_{i+1} - H_i,$$

where  $H_i := H_i(t) = H(t, i/N)$  and  $H$  is a fixed smooth function of  $t$  and  $x$ . For large  $N$ ,  $E_i(t) \approx N^{-1} \partial_x H(t, i/N)$ .

Following similar steps as in the previous section for BEP( $m$ ) but now with the generator

$$L_N^{\text{WABEP}(m)} = L_N^{\text{BEP}(m)} + \Gamma,$$

we have

$$\Gamma \langle \pi_N(Z(t)), \phi \rangle = N \sum_{i=1}^N \phi_i \Gamma z_i,$$

and from the choice of  $\Gamma$  above we obtain

$$\begin{aligned} \Gamma z_i &= -N^2 \left[ E_{i-1} z_{i-1} z_i (\partial_{i-1} - \partial_i) + E_i z_i z_{i+1} (\partial_i - \partial_{i+1}) \right] z_i \\ &= N^2 [E_{i-1} z_{i-1} z_i - E_i z_i z_{i+1}], \end{aligned}$$

so that

$$\Gamma \langle \pi_N(Z(t)), \phi \rangle = N \sum_{i=1}^N \phi_i [E_{i-1} z_{i-1} z_i - E_i z_i z_{i+1}].$$

Again applying summation by parts, we get

$$\Gamma \langle \pi_N(Z(t)), \phi \rangle = N \sum_{i=1}^N (\phi_{i+1} - \phi_i) E_i z_i z_{i+1}.$$

Combining this with the computation in the previous section we find

$$\begin{aligned} L_N^{\text{WABEP}(m)} \langle \pi_N(Z(t)), \phi \rangle &= mN \sum_{i=1}^N z_i [\phi_{i-1} + \phi_{i+1} - 2\phi_i] + N \sum_{i=1}^N (\phi_{i+1} - \phi_i) E_i z_i z_{i+1} \\ &= \frac{m}{N} \sum_{i=1}^N z_i \Delta_N \phi_i + \frac{1}{N} \sum_{i=1}^N z_i z_{i+1} \nabla_N H_i \nabla_N \phi_i \\ &\approx m \langle \pi_N(Z(t)), \partial_{xx} \phi \rangle + \langle [\pi_N^2(Z(t))] \partial_x H, \partial_x \phi \rangle, \end{aligned}$$

where we write  $\nabla_N H_i = N(H_{i+1} - H_i)$  for the discrete derivative and we define

$$\pi_N^{(2)}(Z(t)) := \frac{1}{N} \sum_{i=1}^N z_i z_{i+1} \delta_{i/N}.$$

Note that  $\pi_N^{(2)}$  is not simply the square of  $\pi_N$ ; however, in the limit  $N \rightarrow \infty$  we expect that  $\text{w-lim}_{N \rightarrow \infty} \pi_N^{(2)}(Z) = \left( \text{w-lim}_{N \rightarrow \infty} \pi_N(Z) \right)^2$ , as we will see in the next section. This property is crucial in order to obtain a closed equation for  $\rho(t, x)$ , and is related to the fact that in the hydrodynamic time scale  $N^2$ , local equilibrium installs.

## VII. THE REPLACEMENT LEMMA

We now discuss the Replacement Lemma for the WABEP( $m$ ) in more detail. The term

$$\frac{1}{N} \sum_{i=1}^N z_i z_{i+1} \nabla_N H_i \nabla_N \phi_i \quad (23)$$

is not a function of  $\pi_N(Z)$ , and in order to close the equation we need to replace this term by a function of  $\pi_N(Z)$ . This is a classical problem in the theory of hydrodynamic limits, and various methods have been devised to tackle it<sup>27–29</sup>. In the context of a Ginzburg-Landau model Guo, Papanicolaou, and Varadhan<sup>27,30</sup> show that a term of the form (23) can be replaced by a function of  $\pi_N$ . More precisely, for  $\psi : \mathbb{T} \rightarrow \mathbb{R}$  smooth and any  $\epsilon > 0$ ,

$$\begin{aligned} \frac{1}{N} \sum_{i=1}^N z_i(t) z_{i+1}(t) \psi_i &= \frac{1}{N} \sum_{i=1}^N \frac{1}{2\epsilon N + 1} \sum_{|j-i| \leq \epsilon N} z_j(t) z_{j+1}(t) \psi_i + O(\epsilon) \\ &= \frac{1}{N} \sum_{i=1}^N \nu_{\frac{1}{2\epsilon N + 1}} \sum_{|j-i| \leq \epsilon N} z_j(t) (z_0 z_1) \psi_i + O(\epsilon) \\ &= \frac{1}{N} \sum_{i=1}^N \left( \frac{1}{2\epsilon N + 1} \sum_{|j-i| \leq \epsilon N} z_j(t) \right)^2 \psi_i + O(\epsilon). \end{aligned} \quad (24)$$

The first equality follows from the smoothness of  $\psi$ . In the second step  $\nu_\rho$  is the product measure (12) on the two-site state space with variables  $z_0$  and  $z_1$ , with  $\mathbb{E}_{\nu_\rho}(z_0) = \mathbb{E}_{\nu_\rho}(z_1) = \rho$ ,  $\nu_\rho(f) = \int f d\nu_\rho$ , and hence  $\nu_\rho(z_0 z_1) = \rho^2$ . Given this definition, the third equality follows from these properties of  $\nu_\rho$ . The main approximation step is therefore the second equality. Once one has justified this approximation, the expression in the last line is again a function of the empirical energy profile.



Intuitively, the approximation in the second equality arises from the local-equilibrium property. However, since one wants to apply it in the context of large deviations, one has to show that the approximation is superexponentially good. This means that if we denote for given  $\delta > 0$  the event

$$\left| \int_0^T \left( \frac{1}{N} \sum_{i=1}^N \frac{1}{2\epsilon N + 1} \sum_{|j-i| \leq \epsilon N} z_j(N^2 t) z_{j+1}(N^2 t) \psi_i - \frac{1}{N} \sum_{i=1}^N \nu_{\frac{1}{2\epsilon N + 1}} \sum_{|j-i| \leq \epsilon N} z_j(N^2 t) (z_0 z_1) \psi_i \right) ds \right| > \delta$$

by  $A_\delta^N$ , then

$$\limsup_{N \rightarrow \infty} \frac{1}{N} \log \mathbb{P}_N^{\text{BEP}(m)}(A_\delta^N) = -\infty. \quad (25)$$

For the Ginzburg-Landau model of 27 the superexponential estimate and the dynamical large-deviation principle was proven in 31. See 26 for more details and a proof in the context of gradient zero range processes. To apply the same method to the WABEP there are two additional technical difficulties, similar to the case of the KMP model<sup>32</sup>, namely first that the equilibrium invariant measures have only exponential tails, and second that the mobility is equal to  $\rho^2$  and thus is unbounded.

Assuming that the Replacement Lemma holds, we can proceed with the hydrodynamic limit of the WABEP( $m$ ), and we obtain from the computation in the two previous sections

$$L_N^{\text{WABEP}(m)} \langle \pi_N(Z(t)), \phi \rangle \approx m \int \rho(t, x) \partial_{xx} \phi(x) + \int \rho(t, x)^2 \partial_x H(t, x) \partial_x \phi(x),$$

which leads to the hydrodynamic limit given by the equation

$$\partial_t \rho = m \partial_{xx} \rho - \partial_x (\rho^2 \partial_x H). \quad (26)$$

## VIII. LARGE DEVIATIONS FROM THE HYDRODYNAMIC LIMIT

In the language of the general large-deviation theory developed by 33, equation (19) is a model with diffusion constant  $D = m$  and susceptibility (mobility)  $\chi(\rho) = \rho^2$ . Assuming the validity of the superexponential Replacement Lemma, and as a consequence the limiting equation for the WABEP( $m$ ), the road to large deviations is well known. We describe this procedure here somewhat intuitively. The hydrodynamic limit for the BEP( $m$ ) predicts that  $\pi_N(Z(t))$  evolves “typically” as  $\rho(t, x) dx$  where  $\rho(t, x)$  solves the diffusion equation (19). This behavior is a manifestation of the law of large numbers and therefore one expects

corresponding exponentially small (in  $N$ ) large-deviation probabilities, i.e., we expect, in the sense of large deviations

$$\mathbb{P}_N^{\text{BEP}(m)}(\pi_N(Z)|_{[0,T]} \approx \gamma dx|_{[0,T]}) \sim e^{-NI(\gamma)}. \quad (27)$$

For the rate function  $I$  one expects the Lagrangian form

$$I(\gamma) = \int_0^T \int_{\mathbb{T}} \mathcal{L}(\gamma(t, x), \dot{\gamma}(t, x)) dx dt.$$

In order to find  $\mathcal{L}$ , one modifies the BEP by adding a weak time-dependent asymmetry  $H(t, x)$  as we described above, in such a way that the hydrodynamic equation (26) has the trajectory  $\gamma$  as a solution, i.e., we look for  $H = H(t, x)$  such that

$$\partial_t \gamma(t, x) = m \partial_{xx}(\gamma(t, x)) - \partial_x(\gamma(t, x)^2 \partial_x H(t, x)). \quad (28)$$

Here  $H(t, x)$  is the unknown function, i.e., this is a linear Cauchy problem in  $H$  for given  $\gamma$ . The solution  $H$  can be written formally as

$$\partial_x H(t, x) = \frac{1}{\gamma(t, x)^2} \partial_x^{-1} (\partial_t \gamma(t, x) - m \partial_{xx}(\gamma(t, x))) \quad (29)$$

With this choice of  $H$  the event  $\pi_N(Z) \approx \gamma$  becomes “typical”, i.e., of probability close to one as  $N \rightarrow \infty$ . Therefore, one can use the “Cramér” tilting method:

$$\begin{aligned} \log \mathbb{P}_N^{\text{BEP}(m)}(\pi_N(Z) \approx \gamma dx) &= \log \int \mathbb{1}(\pi_N \approx \gamma dx) \frac{d\mathbb{P}_N^{\text{BEP}(m)}}{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}} d\mathbb{P}_{N,H}^{\text{WABEP}(m)} \\ &\geq \int \mathbb{1}(\pi_N \approx \gamma dx) \log \frac{d\mathbb{P}_N^{\text{BEP}(m)}}{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}} d\mathbb{P}_{N,H}^{\text{WABEP}(m)} \end{aligned}$$

where  $\mathbb{1}$  is the indicator function. So one has to compute

$$\log \frac{d\mathbb{P}_N^{\text{BEP}(m)}}{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}}.$$

The Radon-Nikodym derivative  $d\mathbb{P}_N^{\text{BEP}(m)} / d\mathbb{P}_{N,H}^{\text{WABEP}(m)}$  is given by the Girsanov formula (see the appendix):

$$\frac{1}{N} \log \frac{d\mathbb{P}_N^{\text{BEP}(m)}}{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}} = -\frac{1}{8} \int_0^T \sum_i E_i(t) z_i z_{i+1} + \mathcal{M} \quad (30)$$

with  $E_i(t) = \partial_x H(i/N, t)$ . Here  $\mathcal{M}$  is a martingale under  $\mathbb{P}_{N,H}^{\text{WABEP}(m)}$ . Using the superexponential estimate (25) this is rewritten in terms of the density field, which as in the derivation of the hydrodynamic limit leads to

$$\frac{1}{N} \int \mathbb{1}(\pi_N \approx \gamma dx) \log \frac{d\mathbb{P}_N^{\text{BEP}(m)}}{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}} d\mathbb{P}_{N,H}^{\text{WABEP}(m)} \approx -\frac{1}{8} \int_0^T \int (\partial_x H(x, t))^2 \gamma(t, x)^2 dx dt.$$

Therefore the large deviations of the time courses of  $\pi_N(Z)$  are described by

$$\mathbb{P}_N^{\text{BEP}(m)}(\pi_N(Z)|_{[0,T]} \approx \rho|_{[0,T]}) \sim \exp\{-NI_{\rho^0}^{\text{BEP}(m)}(\rho)\},$$

where the rate function  $I_{\rho^0}^{\text{BEP}(m)}$  is

$$I_{\rho^0}^{\text{BEP}(m)}(\rho) = \begin{cases} \frac{1}{8} \int_0^T \int_{\mathbb{T}} \rho(t, x)^2 (\partial_x H(t, x))^2 dx dt, & \text{if } \rho(0, \cdot) = \rho^0 \text{ and } H \text{ is given by (26),} \\ +\infty & \text{otherwise.} \end{cases}$$

Note that  $I_{\rho^0}^{\text{BEP}(m)}$  is a measure of the degree to which  $\rho$  does *not* satisfy the equation  $\partial_t \rho = m \partial_{xx} \rho$ .

## IX. THE GBEP( $a$ ) AND KMP PROCESSES

The calculations above were done for the Brownian Energy Process with parameter  $m$ . For the GBEP( $a$ ) and KMP processes the arguments are very similar. In fact, the invariant measures of the GBEP( $a$ ) are the same as for BEP(1), and for the KMP the same as for BEP(2). This also implies that the large-deviation rate functionals for the two processes are  $S_{\rho_0,1}$  and  $S_{\rho_0,2}$ .

Therefore the three processes have equilibrium rate functionals  $S_{\rho_0,m}$ ,  $S_{\rho_0,1}$ , and  $S_{\rho_0,2}$  that are essentially the same; they only differ by a multiplicative constant.

For the GBEP( $a$ ), the addition of the factor  $a$  in (9) and (10) modifies the hydrodynamic limit, which now becomes

$$\partial_t \rho = \partial_x (a^2(\rho) \partial_x \rho).$$

For the large deviations of the time courses of the empirical measure  $\pi_N$ , similar arguments as for BEP( $m$ ) give the expression

$$I_{\rho^0}^{\text{GBEP}(a)}(\rho) = \frac{1}{8} \int_0^T \int_{\mathbb{T}} \rho^2 a^2(\rho) (\partial_x H)^2, \quad \partial_t \rho - \partial_x (a^2(\rho) \partial_x \rho) = \partial_x (\rho^2 a^2(\rho) \partial_x H), \quad (31)$$

provided that  $\rho(0, \cdot) = \rho^0$ .

For the KMP, the hydrodynamic limit and the large deviations were found in 32 to be

$$\partial_t \rho = \partial_{xx} \rho$$

and

$$I_{\rho^0}^{\text{KMP}}(\rho) = \frac{1}{4} \int_0^T \int_{\mathbb{T}} \rho^2 (\partial_x H)^2, \quad \partial_t \rho - \partial_{xx} \rho = \partial_x (\rho^2 \partial_x H), \quad (32)$$

again provided that  $\rho(0, \cdot) = \rho^0$ .

## X. GRADIENT-FLOW STRUCTURES

We now describe how these results identify gradient-flow structures for the limiting equations. Above we have derived three large-deviation rate functionals  $I_{\rho^0}^{\text{BEP}(m)}$ ,  $I_{\rho^0}^{\text{GBEP}(a)}$ , and  $I_{\rho^0}^{\text{KMP}}$ , which each can be written in the form

$$I(\rho) := \frac{1}{2} \int_0^T \left\| \partial_t \rho - K_\rho \frac{\delta E}{\delta \rho} \right\|_{K_\rho^{-1}}^2 dt, \quad \text{provided } \rho(0, \cdot) = \rho^0, \quad (33)$$

where

$$E(\rho) := -\gamma \int_{\mathbb{T}} \log \rho \quad \text{and} \quad K_\rho \xi := -\partial_x (\alpha(\rho) \partial_x \xi). \quad (34)$$

Here  $\gamma$  and  $\alpha$  are process-dependent. We list their values and the corresponding limiting equation:

$$\text{BEP}(m) : \quad \gamma = m/4 \quad \text{and} \quad \alpha(\rho) = 4\rho^2; \quad \partial_t \rho = m \partial_{xx} \rho; \quad (35a)$$

$$\text{GBEP}(a) : \quad \gamma = 1/4 \quad \text{and} \quad \alpha(\rho) = 4\rho^2 a(\rho)^2; \quad \partial_t \rho = \partial_x (a^2(\rho) \partial_x \rho); \quad (35b)$$

$$\text{KMP} : \quad \gamma = 1/2 \quad \text{and} \quad \alpha(\rho) = 2\rho^2; \quad \partial_t \rho = \partial_{xx} \rho. \quad (35c)$$

Note that for each of the processes  $E$  is equal to one-half of the corresponding equilibrium rate functional  $S_{\rho_0, m}$ ,  $S_{\rho_0, 1}$ , or  $S_{\rho_0, 2}$  up to an additive constant, since  $\int \rho$  can be assumed constant during evolution.

The norm  $\|\cdot\|_{K_\rho^{-1}}$  in (33) is defined as the dual norm of

$$\|\xi\|_{K_\rho}^2 := \int_{\mathbb{T}} \xi K_\rho \xi,$$

Here, for a norm  $\|\cdot\|$  on a linear space  $X$ , the dual norm  $\|\cdot\|_*$  on the dual space  $X'$  is defined via

$$\|f\|_*^2 = \sup \frac{|\langle f, g \rangle|^2}{\|g\|^2}$$

where  $\langle f, g \rangle$  denotes the pairing between  $X'$  and  $X$ . It can be written in terms of a formal inverse  $G_\rho = K_\rho^{-1}$  as

$$\|s\|_{K_\rho^{-1}}^2 = \|s\|_{G_\rho}^2 := \int_{\mathbb{T}} s G_\rho s.$$

To formally verify the claim above that  $I_{\rho^0}^{\text{BEP}(m)}$  can be written in the form (33), for instance, first note that  $K_\rho \delta E / \delta \rho = m \partial_{xx} \rho$ . Then, whenever  $\partial_t \rho - m \partial_{xx} \rho = \partial_x (\rho^2 \partial_x H)$ ,

$$K_\rho H = -4(\partial_t \rho - m \partial_{xx} \rho) \quad \text{so that} \quad H = -4G_\rho (\partial_t \rho - m \partial_{xx} \rho).$$

Therefore

$$\begin{aligned} I_{\rho^0}^{\text{BEP}(m)}(\rho) &= \frac{1}{8} \int_0^T \int_{\mathbb{T}} \rho^2 (\partial_x H)^2 = -\frac{1}{8} \int_0^T \int_{\mathbb{T}} H \partial_x (\rho^2 \partial_x H) \\ &= \frac{1}{32} \int_0^T \int_{\mathbb{T}} H K_{\rho} H = \frac{1}{2} \int_0^T \int_{\mathbb{T}} (\partial_t \rho - m \partial_{xx} \rho) G_{\rho} (\partial_t \rho - m \partial_{xx} \rho) = I(\rho). \end{aligned}$$

The structure (33) identifies a gradient-flow structure with state space  $L_{\geq 0}^1(\mathbb{T})$ , with driving functional  $E$ , and with Onsager operator  $K_{\rho}$ . This is the structure that we are looking for.

## XI. CONCLUSIONS AND DISCUSSION

In the Introduction we asked the question ‘what is the natural gradient-flow structure for the conduction of heat?’ By considering three different microscopic stochastic processes that model heat conduction we identified, through their large deviations, three gradient structures for macroscopic heat equations. We now comment on our findings.

### The driving functional $E$

Each of the three gradient-flow structures (35) is driven by a functional  $E$ , which is the same, up to multiplicative and additive constants, as the large-deviation rate functional  $S_{\rho}$ , of the equilibrium invariant measure of the process. Therefore, despite the differences in the *dynamics* of the processes, the functional that *drives* those dynamics is nearly the same for the three processes. This is of course a consequence of the similarity of the invariant measures themselves, and therefore it is interesting to understand these similarities.

Although it is difficult to pin down necessary and sufficient conditions leading to the invariant measures of these processes, one can identify some ingredients:

- The variable  $p_i$  (or the variables  $p_i^j$ ) represents a quantity on a *fixed* lattice;
- The evolution locally preserves the quantity  $p_i^2 + p_{i+1}^2$  (or  $\sum_{j,k=1}^m (p_i^j)^2 + (p_{i+1}^k)^2$ );

In the thermodynamic limit ( $N \rightarrow \infty$ ), using the equivalence of ensembles results in an invariant product measure with Lebesgue density  $\exp(-e_i/\theta)$ , at each site  $i$  where  $e_i = p_i^2$  or  $e_i = \sum_j (p_i^j)^2$ . After transition to the variables  $z_i$ , this invariant measure transforms into

the Gamma distribution  $\nu_{N,\theta,m}$  defined in (12); all Gamma distributions lead to the same form of  $E$ , with  $-\gamma \log \rho$  as the important term under the integral.

Summarizing, the form of the driving functional  $E$  is related to the occurrence of a conserved quadratic energy  $p^2$  on a fixed lattice.

There is also a classical thermodynamical argument that leads to a driving functional proportional to the logarithm of the temperature. We start from the postulate for simple closed homogeneous systems without volume change,

$$TdS = dQ \quad (36)$$

where  $dS$  is the infinitesimal change in entropy  $S$  while adding  $dQ$  of heat to a system at temperature  $T$  (the barred  $d$  is the usual way of indicating that  $dQ$  is not necessarily an exact differential (see e.g. 34, Ch. 1)). Taking the example of BEP( $m$ ), internal energy is proportional to temperature (see the relationship between  $\theta$  and  $\mathbb{E}_{\nu_{\theta,m}}[z_i]$  in (13)):  $U = mT/2$ . Therefore  $dQ$  is equal to the change  $dU$  in internal energy  $U$ , and (36) reduces to

$$dS = \frac{m}{2} \frac{1}{U} dU = \frac{m}{2} d \log U.$$

If we accept the interpretation of the thermodynamic entropy  $S$  as (minus) the large-deviation rate functional of the equilibrium system, then this relation provides a separate argument why the driving functional should be proportional to the logarithm of the temperature.

## The Onsager operator $K$

While  $S_{\rho,\cdot}$  and therefore  $E$  is generated by the *conservation* properties of the process, the rates in the generator determine the ‘diffusion constant’  $\alpha(\rho)$  in  $K_{\rho}$  (34). The example of the GBEP( $a$ ) (35b) shows that there is nearly full freedom here:  $\alpha(\rho) = a^2(\rho)$  can be any positive function of  $\rho$ , and under certain conditions this coefficient can even vanish at certain values of  $\rho$  (‘degenerate diffusion’ 35).

A special feature of the BEP( $m$ ) and KMP models is that the mobility is quadratic ( $\alpha(\rho) \sim \rho^2$ ) and the driving functional logarithmic; the two combine to provide a *linear* heat equation, through  $\partial_x(\rho^2 \partial_x(-1/\rho)) = \partial_{xx}\rho$ . It is interesting to ask where this quadratic mobility comes from.

In the BEP( $m$ ) the quadratic mobility can be traced back to two ingredients:

1. The stochastic force that lattice point  $i + 1$  exerts on its neighbor at  $i$  scales as  $p_{i+1}$  (see the first two terms in (5)), and therefore its effect in  $L^{\text{BMP}}$  as  $p_{i+1}^2$  (see (4)).
2. Since  $dp_i^2/dt = 2p_i dp_i/dt$ , the effect of this forcing on  $z_i$  is multiplied by  $p_i$  in the SDE (7) and by  $p_i^2$  in  $L^{\text{BEP}(m)}$  (see (8)).

These are the reasons for the mobility  $z_i z_{i+1}$  that multiplies the second-order term in (8).

For the KMP model, there is an additional explanation. Since transfer of energy happens by redistributing energy with a uniform distribution, the expected transfer of energy scales linearly with the difference of the expected energies in the two lattice points. In formulas,

$$L_{i,i+1}^{\text{KMP}} z_i = \int_0^1 [s(z_i + z_{i+1}) - z_i] ds = \frac{1}{2}(z_i + z_{i+1}).$$

This linear energy flux translates into a linear hydrodynamic limit.

## Comparison with diffusion

Comparing these systems with diffusion systems we observe some similarities and some differences.

*Similarity 1:* Both concern the redistribution over time of a conserved quantity: energy for heat conduction and mass for diffusion. Consequently both have hydrodynamic limits in divergence form.

*Difference 1:* The stationary rate functionals are different:  $-\int \log \rho$  for heat conduction, versus  $\int \rho \log \rho$  for diffusion. These differences reflect the different origins:  $-\int \log \rho$  arises from a Gamma-distributed quantity *at* each lattice point, while  $\int \rho \log \rho$  arises from the *moving-around* of masses, through Stirling's formula (see e.g. 36, Sec. 4.2).

*Difference 2:* The mobilities are different:  $\rho^2$  for KMP and  $\text{BEP}(m)$ ,  $\rho^2 a^2(\rho)$  for  $\text{GBEP}(a)$ , whereas diffusion has mobility  $\rho$ . Again this reflects the different origins: as described above, the mobility  $\rho^2$  for KMP and  $\text{BEP}(m)$  arises from the combination of neighboring forces that scale as  $p$  with energies that scale as  $p^2$ , while the mobility  $\rho$  for diffusion arises from a simple counting argument, where the flux of particles is proportional to the number of particles.

In this context it is instructive also to compare to the Simple Symmetric Exclusion Process, the jump process for particles on a lattice in which particles jump left and right

with rate 1 whenever the destination site is empty. Here the mobility is  $\rho(1 - \rho)$ , which can be constructed through the same counting argument as above (the factor  $\rho$ ) mitigated by the probability of finding the neighboring site empty (the factor  $(1 - \rho)$ ).

## No metric space

When the mobility  $\alpha(\rho)$  in  $K_\rho$  is concave, a generalized Wasserstein-type metric can be constructed as an infimum along curves, following 37:

$$d(\rho_0, \rho_1)^2 := \inf_{\rho, w} \left\{ \int_0^1 \int_{\mathbb{T}} \frac{w^2}{\alpha(\rho)} : \partial_t \rho = \partial_x w \text{ and } \rho|_{t=0,1} = \rho_{0,1} \right\}. \quad (37)$$

The concaveness of  $\alpha$  causes the function  $(w, r) \mapsto w^2/\alpha(r)$  to be convex, rendering the integral above lower-semicontinuous with respect to various weak topologies for  $w$  and  $\rho$ . This property is used in 38 to construct gradient-flow solutions with respect to this metric, in the sense of 39, obtaining various other properties in the process.

However, if  $\alpha$  is *not* concave, then this metric need not be well-defined; for instance, if  $\lim_{r \rightarrow \infty} \alpha(r)/r = +\infty$ , then it is not difficult to see that for each  $\rho_0, \rho_1$  with  $\int(\rho_0 - \rho_1) = 0$  the infimum in (37) is zero<sup>40</sup>. The important case  $\alpha(\rho) = \rho^2$  (BEP( $m$ ) and KMP) is an example of this; the natural metric (37) is therefore not well defined for these systems.

## Generalizations to other models

Although the fundamental question of this paper—what is the natural gradient structure for heat conduction?—is most simply posed for the simplest of heat-conducting systems, the insights that we gained are also useful for more complex models. For instance, in the ‘variational-modeling’ philosophy<sup>36</sup> one constructs mathematical models of real-world systems by choosing a driving functional and an Onsager operator. These two choices fully determine the system. For this methodology to function well, one needs to understand how to choose these components—and this is exactly the question of this paper.

## Acknowledgements

The authors are grateful for many interesting discussions with Jin Feng. MAP and KV acknowledge the support of NWO VICI grant 639.033.008.



## Appendix A: A Girsanov Formula

Here we prove (30). We refer to 41 for more details on the proof of the Girsanov formula in the context of diffusion processes; the formula is proved for bounded coefficients in 41, Th. 6.4.2, and the extension to unbounded coefficients can be done with the methods of 41, Ch. 11.

Consider a Markov process with generator

$$L = \frac{1}{2} \sum_i \alpha^i(x) \partial_i^2 + \sum_i \beta^i(x) \partial_i \quad (\text{A1})$$

and another Markov process obtained from  $L$  by addition of an extra drift term of the form

$$\hat{L} = \frac{1}{2} \sum_i \alpha^i(x) \partial_i^2 + \sum_i \beta^i(x) \partial_i + \sum_i \alpha^i(x) \gamma_t^i(x) \partial_i. \quad (\text{A2})$$

These two processes correspond to the following stochastic differential equations:

$$dx_t^i = \beta^i(x_t) dt + \sqrt{\alpha^i(x_t)} dB_t^i \quad (\text{A3})$$

and

$$d\hat{x}_t^i = (\beta^i(x_t) + \alpha^i(x_t) \gamma_t^i(x_t)) dt + \sqrt{\alpha^i(x_t)} dB_t^i. \quad (\text{A4})$$

Let us call  $\mathbb{P}$ , resp.  $\hat{\mathbb{P}}$  the path space measure of the  $x$ , resp.  $\hat{x}$  process. The Girsanov formula then gives the Radon Nikodym derivative of the  $\hat{x}$  process with respect to the  $x$  process:

$$\frac{d\hat{\mathbb{P}}}{d\mathbb{P}} = \exp \left( \sum_i \left( \int_0^T \gamma_t^i(x_t) (dx_t^i - \beta^i(x_t) dt) - \frac{1}{2} \int_0^T (\gamma_t^i)^2(x_t) \alpha^i(x_t) dt \right) \right). \quad (\text{A5})$$

As a consequence, in the case  $\gamma_t^i = \gamma(\frac{i}{N}, t)$ , we find

$$\begin{aligned} \frac{1}{N} \log \frac{d\hat{\mathbb{P}}}{d\mathbb{P}} &= \frac{1}{N} \sum_i \left( \int_0^T \gamma_t^i(x_t) (dx_t^i - \beta^i(x_t) dt) - \frac{1}{2} \int_0^T (\gamma_t^i)^2(x_t) \alpha^i(x_t) dt \right) \\ &= \frac{1}{2N} \sum_i \int_0^T (\gamma_t^i)^2(x_t) \alpha^i(x_t) dt + \mathcal{M}, \end{aligned} \quad (\text{A6})$$

where

$$\mathcal{M} := \frac{1}{N} \sum_i \int_0^T \gamma_t^i(x_t) [dx_t^i - \beta^i(x_t) dt - \alpha^i(x_t) \gamma_t^i(x_t) dt]$$

is a martingale under  $\hat{\mathbb{P}}$ . We can now apply this to the case of Radon-Nikodym derivative of the BEP( $m$ ) with respect to the WABEP( $m$ ):

$$L_N^{\text{BEP}(m)} = \sum_i 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 - m(z_i - z_{i+1}) (\partial_i - \partial_{i+1})$$

and

$$L_{N,H}^{\text{WABEP}(m)} = \sum_i 2z_i z_{i+1} (\partial_i - \partial_{i+1})^2 - m(z_i - z_{i+1}) (\partial_i - \partial_{i+1}) - E_i z_i z_{i+1} (\partial_i - \partial_{i+1}).$$

We make the change of variables to  $y_i := z_i - z_{i+1}$ , and defining  $\alpha_i(y) = 4z_i z_{i+1}$ ,  $\beta_i(y) = -m y_i$  and  $\gamma_i(y) = -\frac{1}{4} E_i = -\frac{1}{4} E(\frac{i}{N}, u)$  we arrive at the same situation as before to obtain

$$\frac{1}{N} \log \frac{d\mathbb{P}_{N,H}^{\text{WABEP}(m)}}{d\mathbb{P}_N^{\text{BEP}(m)}} = \frac{1}{8} \int_0^T \sum_i E_i^2 z_i z_{i+1} dt + \mathcal{M},$$

where  $\mathcal{M}$  is a martingale under  $\mathbb{P}_{N,H}^{\text{WABEP}(m)}$ .

## REFERENCES

- <sup>1</sup>S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. From a large-deviations principle to the Wasserstein gradient flow: A new micro-macro passage. *Communications in Mathematical Physics*, 307:791–815, 2011.
- <sup>2</sup>S. Adams, N. Dirr, M. A. Peletier, and J. Zimmer. Large deviations and gradient flows. *Arxiv preprint arXiv:1201.4601*, 2012. Submitted.
- <sup>3</sup>M. H. Duong, V. Laschos, and D. R. M. Renger. Wasserstein gradient flows from large deviations of many-particle limits. *ESAIM: Control, Optimisation and Calculus of Variations*, E-first, 2013.
- <sup>4</sup>M. A. Peletier, D. R. M. Renger, and M. Veneroni. Variational formulation of the Fokker-Planck equation with decay: A particle approach. *Communications in Contemporary Mathematics*, 15(5):1350017, 2013.
- <sup>5</sup>M. H. Duong, M. A. Peletier, and J. Zimmer. Conservative-dissipative approximation schemes for a generalized Kramers equation. *arXiv preprint arXiv:1206.2859*, 2012. Submitted.
- <sup>6</sup>M. H. Duong, M. A. Peletier, and J. Zimmer. GENERIC formalism of a Vlasov-Fokker-Planck equation and connection to large-deviation principles. *Nonlinearity*, 26(2951-2971), 2013.
- <sup>7</sup>D. R. M. Renger. *Microscopic Interpretation of Wasserstein Gradient Flows*. PhD thesis, Technische Universiteit Eindhoven, 2013.
- <sup>8</sup>Alexander Mielke, D. R. M Renger, and M. A. Peletier. On the relation between gradient

- flows and the large-deviation principle, with applications to Markov chains and diffusion. *arXiv preprint arXiv:1312.7591*, 2013.
- <sup>9</sup>F. Bonetto, J. L. Lebowitz, and L. Rey-Bellet. Fourier’s law: A challenge to theorists. In A. Fokas, A. Grigoryan, T. Kibble, and B. Zegarlinski, editors, *Mathematical Physics 2000*, pages 128–150. Imperial College Press, London, 2000.
- <sup>10</sup>Abhishek Dhar. Heat transport in low-dimensional systems. *Advances in Physics*, 57(5): 457–537, 2008.
- <sup>11</sup>C Kipnis, C Marchioro, and E Presutti. Heat flow in an exactly solvable model. *Journal of Statistical Physics*, 27(1):65–74, 1982.
- <sup>12</sup>C. Bernardin and S. Olla. Fourier’s law for a microscopic model of heat conduction. *Journal of Statistical Physics*, 121(3-4):271–289, 2005.
- <sup>13</sup>G. Basile, C. Bernardin, and S. Olla. Momentum conserving model with anomalous thermal conductivity in low dimensional systems. *Physical review letters*, 96(20):204303, 2006.
- <sup>14</sup>C. Bernardin. Hydrodynamics for a system of harmonic oscillators perturbed by a conservative noise. *Stochastic processes and their applications*, 117(4):487–513, 2007.
- <sup>15</sup>J. Bricmont and A. Kupiainen. Towards a derivation of Fourier’s law for coupled anharmonic oscillators. *Communications in mathematical physics*, 274(3):555–626, 2007.
- <sup>16</sup>C. Bernardin. Stationary nonequilibrium properties for a heat conduction model. *Physical Review E*, 78(2):021134, 2008.
- <sup>17</sup>G. Basile, C. Bernardin, and S. Olla. Thermal conductivity for a momentum conservative model. *Communications in Mathematical Physics*, 287(1):67–98, 2009.
- <sup>18</sup>C. Bernardin and S. Olla. Transport properties of a chain of anharmonic oscillators with random flip of velocities. *Journal of Statistical Physics*, 145(5):1224–1255, 2011.
- <sup>19</sup>Cristian Giardiná, Jorge Kurchan, and Frank Redig. Duality and exact correlations for a model of heat conduction. *Journal of mathematical physics*, 48:033301, 2007.
- <sup>20</sup>C. Giardiná, J. Kurchan, F. Redig, and K. Vafayi. Duality and hidden symmetries in interacting particle systems. *Journal of Statistical Physics*, 135(1):25–55, 2009.
- <sup>21</sup>C. Giardiná, F. Redig, and K. Vafayi. Correlation inequalities for interacting particle systems with duality. *Journal of Statistical Physics*, 141(2):242–263, 2010.
- <sup>22</sup>F. Redig and K. Vafayi. Weak coupling limits in a stochastic model of heat conduction. *Journal of Mathematical Physics*, 52:093303, 2011.
- <sup>23</sup>Stefan Grosskinsky, Frank Redig, and Kiamars Vafayi. Condensation in the inclusion

- process and related models. *J. Stat. Phys.*, 142(5):952–974, 2011. ISSN 0022-4715.
- <sup>24</sup>This generator is twice the one in [11](#) and [32](#), thus accelerating the evolution by a factor 2.
- <sup>25</sup>A. Dembo and O. Zeitouni. *Large deviations techniques and applications*. Springer Verlag, 1998.
- <sup>26</sup>C. Kipnis and C. Landim. *Scaling limits of interacting particle systems*. Springer Verlag, 1999.
- <sup>27</sup>M. Z. Guo, G. C. Papanicolaou, and S. R. S. Varadhan. Nonlinear diffusion limit for a system with nearest neighbor interactions. *Comm. Math. Phys.*, Volume 118, Number 1: 31–59, 1988.
- <sup>28</sup>H.-T. Yau. Relative entropy and hydrodynamics of Ginzburg-Landau models. *Letters in Mathematical Physics*, 22(1):63–80, 1991. ISSN 0377-9017.
- <sup>29</sup>Natalie Grunewald, Felix Otto, Cédric Villani, and Maria G. Westdickenberg. A two-scale approach to logarithmic Sobolev inequalities and the hydrodynamic limit. *Ann. Inst. Henri Poincaré Probab. Stat.*, 45(2):302–351, 2009. ISSN 0246-0203.
- <sup>30</sup>C. Kipnis, S. Olla, and S. R. S. Varadhan. Hydrodynamics and large deviations for simple exclusion processes. *Commun. Pure Appl. Math.*, 42:115–137, 1989.
- <sup>31</sup>M.D. Donsker and S. R. S. Varadhan. Large deviations from a hydrodynamic scaling limit. *Comm. Pure Appl. Math.*, 42:243–270, 1989.
- <sup>32</sup>L. Bertini, D. Gabrielli, and J. L. Lebowitz. Large deviations for a stochastic model of heat flow. *Journal of Statistical Physics*, 121(5):843–885, 2005.
- <sup>33</sup>L. Bertini, A. De Sole, D. Gabrielli, G. Jona-Lasinio, and C. Landim. Towards a nonequilibrium thermodynamics: A self-contained macroscopic description of driven diffusive systems. *J. Stat. Phys.*, 135(5-6):857–872, 2009. ISSN 0022-4715.
- <sup>34</sup>H. B. Callen. *Thermodynamics and an Introduction to Thermostatistics*. Wiley, 1985.
- <sup>35</sup>M. A. Peletier. *Problems in Degenerate Diffusion*. PhD thesis, University of Leiden, 1997.
- <sup>36</sup>M. A Peletier. Variational modelling: Energies, gradient flows, and large deviations. *Arxiv preprint arXiv:1402.1990*, 2014.
- <sup>37</sup>J.-D. Benamou and Y. Brenier. A computational fluid mechanics solution to the Monge-Kantorovich mass transfer problem. *Numerical Mathematics*, 84:375–393, 2000.
- <sup>38</sup>J.A. Carrillo, S. Lisini, G. Savaré, and D. Slepcev. Nonlinear mobility continuity equations and generalized displacement convexity. *Journal of Functional Analysis*, 258(4):1273–1309, 2010. ISSN 0022-1236.

<sup>39</sup>L. Ambrosio, N. Gigli, and G. Savaré. *Gradient Flows in Metric Spaces and in the Space of Probability Measures*. Lectures in mathematics ETH Zürich. Birkhäuser, 2005.

<sup>40</sup>For this one can observe that transporting mass is vanishingly ‘cheap’ as  $\rho(x) \rightarrow \infty$ , implying that Dirac delta functions can be transported with zero cost. Since  $\rho_0$  and  $\rho_1$  can be transported to purely atomic measures with arbitrarily small cost, it follows that the infimum is zero.

<sup>41</sup>Daniel W. Stroock and S. R. Srinivasa Varadhan. *Multidimensional diffusion processes*, volume 233 of *Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences]*. Springer-Verlag, Berlin, 1979. ISBN 3-540-90353-4.